Minimizing the Reprojection Error in Surface Reconstruction from Images

Pau Gargallo, Emmanuel Prados, Peter Sturm

Minimizing the Reprojection Error in Surface Reconstruction from Images

Pau Gargallo Emmanuel Prados Peter Sturm
INRIA Rhône-Alpes, France
name.surname@inrialpes.fr

Abstract

This paper addresses the problem of image-based surface reconstruction. The main contribution is the computation of the exact derivative of the reprojection error functional. This allows its rigorous minimization via gradient descent surface evolution. The main difficulty has been to correctly take into account the visibility changes that occur when the surface moves. A geometric and analytical study of these changes is presented and used for the computation of derivative.

Our analysis shows the strong influence that the movement of the contour generators has on the reprojection error. As a consequence, during the proper minimization of the reprojection error, the contour generators of the surface are automatically moved to their correct location in the images. Therefore, current methods adding additional silhouettes or apparent contour constraints to ensure this alignment can now be understood and justified by a single criterion: the reprojection error.

1. Introduction

Reconstructing scene models from images is the problem of inverting the image formation process. Many scenes can be well represented by a surface and some additional quantities describing, for example, reflectance properties, lighting conditions or sensor parameters. Such a model, allows to precisely describe how the images were generated from the surface. However, recovering the surface from the images is an old, incompletely solved, computer vision challenge.

In such a context, a solution to the problem would be a surface Γ such that the images generated from the model are most similar to the observed images (i.e. the data). This naturally yields to formulating the problem as the minimization of an error measure between the observed and predicted values of pixels, carried out, importantly, over all pixels in all input images. This is not a trivial task, as will be shown in the following.

For many image formation models, the predicted value of a pixel $\mathbf{u}$ depends only on the position of the point $\pi_\Gamma^{-1}(\mathbf{u})$ that is viewed in that pixel and possibly on its normal $\mathbf{n}(\pi_\Gamma^{-1}(\mathbf{u}))$. This point is the first collision between the viewing ray of $\mathbf{u}$ and the surface Γ or, if non existing, a point in the background B. The error measure between a predicted and an observed image is then of the form

$$E(\Gamma) = \int_{\mathcal{I}} g\big(\, \pi_\Gamma^{-1}(\mathbf{u}),\, \mathbf{n}(\pi_\Gamma^{-1}(\mathbf{u}))\,\big)\, d\mathbf{u}\,, \tag{1}$$

where $\mathcal{I}$ is the set of all pixels in the image, $d\mathbf{u}$ is the area measure on the sensor's image plane, and[1] $g : \mathbb{R}^3 \times \mathbb{S}^2 \to \mathbb{R}$ gives the error measure for the pixel $\mathbf{u}$. We call (1) the *reprojection error functional* and the objective of this paper is to find a method for minimizing it.

This functional class (1) is wide enough to cover many image-based surface reconstruction problems. In section 6, we illustrate an example application to multi-view stereo, where g measures the difference between the observed color of a pixel and the one predicted by the reconstruction. Another example would be the reconstruction from noisy range images [24], where g would measure the difference between the captured depth at $\mathbf{u}$ and the depth of $\pi_\Gamma^{-1}(\mathbf{u})$.

In the last years, great advances on the minimization of surface functionals have been made. Several works have addressed the minimization of the *weighted area functionals*. These are functionals of the form

$$A(\Gamma) = \int_\Gamma g(\mathbf{x}, \mathbf{n}(\mathbf{x}))\, d\sigma\,, \tag{2}$$

where g is integrated on the surface and $d\sigma$ is the surface's area measure. The derivative of this functional has been found, allowing therefore its minimization via gradient descent surface evolution [5, 7, 20]. It has also been shown how to find the global minimum of some of these functionals via graph cuts [2, 11] and continuous max-flow [1].

The difference between the functionals (1) and (2), emanates from the fact that the first is an integral over the image domain, i.e. where the data lives, while the latter is an integral over the surface.

[1] $\mathbb{S}^2$ represents the unit sphere, i.e. the space of normals.

To benefit from the existing knowledge about the weighted area functional, one may try to rewrite the functional (1) as an integral over the surface by counting only the visible points [15, 19, 25]. This gives,

$$E(\Gamma) = -\int_{\Gamma \cup B} g(\mathbf{x}, \mathbf{n}(\mathbf{x}))\, \frac{\mathbf{x} \cdot \mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3}\, \nu_\Gamma(\mathbf{x})\, d\sigma\,, \quad (3)$$

where ν_Γ is the visibility function (giving 1 for an $\mathbf{x}$ that is visible and 0 otherwise, cf. section 4) and where the fact that $d\mathbf{u} = -\frac{\mathbf{x} \cdot \mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3}\, \nu_\Gamma(\mathbf{x})\, d\sigma$ has been used. In order to count all the pixels in the image, the integral extends over the surface but also over the background surface B. This is assumed to be a distant, fixed surface whose projection covers the whole image and whose shape is irrelevant.

We observe that the integrand obtained by the conversion depends on $\mathbf{x}$ and $\mathbf{n}(\mathbf{x})$ as in (2), but also especially on the whole surface Γ, because of the visibility term. Hence, the reprojection error functional is not a weighted area functional and the existing methods for minimizing the weighted area functionals can not be applied.

The main contribution of this paper is the computation of the derivative of the reprojection error functional (section 5), allowing therefore its minimization via gradient descent. To do so, we first study the changes of visibility while a surface moves (section 4). We will particularly observe that contour generators have a strong influence on these changes. When a contour generator moves, some hidden parts of the surface or the background appear behind it and some visible parts disappear. The backprojection $\pi_\Gamma^{-1}(\mathbf{u})$ of the pixels at the corresponding apparent contour moves suddenly from one part of the surface to another. This has a strong effect on the predicted value of these pixels and therefore on the reprojection error and its derivative.

As a consequence, the correct gradient descent evolution of the reprojection error automatically favors and ensures the alignment of the apparent contours of the reconstructed surface with discontinuities present in the images. This alignment thus provides a generalization of the visual hull that takes into account all the apparent contours and not only the silhouettes (i.e. *outer* apparent contours). The experiments of section 6 will demonstrate this alignment in a particular application of the functional to multi-view stereo.

2. Related Work

Most state of the art surface reconstruction algorithms [17] use, at some point, a weighted area functional. The cost of a surface point is defined by a photo-consistency measure using the images where this point is visible. Not being possible to include the visibility in the functional itself, it has to be determined before evolving the surface. This can be done once and for all [8, 14, 23] or iteratively, alternating the computation of the visibility with the optimization of the functional [5, 15, 21].

Any method not including the visibility in the functional suffers, to some extend, of the *minimal surface bias* [1, 26]. This is a bias towards small surfaces. Its most notable effect is that the null surface has cost 0 and is therefore the global minimum. A softer effect is the tendancy of small and thin parts of the surface to disappear.

Palliatives have been proposed. Ballooning forces [23] pump the surface to avoid shrinkage and tend to get balloon like results [26]. Surface evolution methods [5, 15] rely implicitly on the fact that, for sufficiently textured surfaces, a wide local minimum exists close to a good reconstruction. Thus, the evolution will stop before shrinking too much. Visual hull based approaches constrain the surface to fill the silhouettes of the object in the images [8, 6, 18]; the bias is thus reduced, but only in parts of the surface that are close to the visual hull.

Stereoscopic segmentation [25] use the concept of oriented visibility [11] to include the visibility in a weighted area functional. In consequence, the shrinkage is avoided and the resulting surface is consistent with the silhouettes in the images. This happens automatically without the need of additional constraints. However, the oriented visibility approximation is only valid for *convex* objects and the evolution derived in [25] does not correctly handle self-occlusions.

Visual hull constraints have been generalized to taking into account not only silhouettes but all apparent contours generators, by enforcing them to be aligned with strong image gradients [4, 9]. The same way that the stereoscopic segmentation manage to reconstruct visual hull like surfaces without silhouettes constraints, the proper minimization of the reprojection error presented in this paper performs the aligment of all the apparent contours naturally, without any additional constaints.

3. Mathematical Background and Notation

The mathematical framework used in this paper is the one defined by Solem and Overgaard [20] in which shapes are implicitly represented by level set functions [12, 13]. For the convenience of the reader, we remind the related notions and notations required for understanding our work.

3.1. Level Set and Characteristic Functions

Given a level set function $\phi : \mathbb{R}^3 \to \mathbb{R}$, the set of points $\Omega = \{\mathbf{x} : \phi(\mathbf{x}) \leq 0\}$ is a *solid shape* [10] and its boundary $\Gamma = \{\mathbf{x} : \phi(\mathbf{x}) = 0\}$ is an *oriented surface*. We say that ϕ is an *implicit representation* of Ω and that Ω is the *inside* of Γ. The outward normal vector of the surface can be computed from the implicit representation as $\mathbf{n} = \nabla\phi / |\nabla\phi|$.

The *characteristic function* of the shape, χ_Ω, evaluates to 1 inside the shape and 0 outside. It can easily be expressed in terms of ϕ and the Heaviside step function, H, as $\chi_\Omega = 1 - H(\phi)$. The gradient $\nabla\chi_\Omega$ of the characteristic function can only be defined in the distributional sense [16]. For all test vector fields $\mathbf{w} : \mathbb{R}^3 \to \mathbb{R}^3$,

$$\int_{\mathbb{R}^3} \nabla \chi_\Omega \cdot \mathbf{w} \, d\mathbf{x} \equiv -\int_\Omega \nabla \cdot \mathbf{w} \, d\mathbf{x} = -\int_\Gamma \mathbf{w} \cdot \mathbf{n} \, d\sigma \,, \quad (4)$$

where the last term results from Gauss' divergence theorem. In other words, the distribution $\nabla \chi_\Omega$ computes the flux of $\mathbf{w}$ that is entering the shape. An expression in terms of ϕ can be obtained by the chain rule:

$$\nabla \chi_\Omega = -\nabla \phi \, \delta(\phi) \,, \quad (5)$$

where δ is the Dirac delta distribution.

3.2. Functional Derivatives

Let M denote the *manifold of admissible surfaces* defined by Solem and Overgaard [20]. Points in this space are surfaces. The tangent vectors on a point Γ are the normal velocities by which the surface can evolve. The tangent space $T_\Gamma M$ is the set of all these normal velocities.

The variation $\phi^s = \phi + s\psi$ of ϕ, describes a curve $\Gamma(s) = \{\mathbf{x} : \phi^s(\mathbf{x}) = 0\}$ in M. Its *normal velocity* (or tangent vector) at $s = 0$ is,

$$v = \frac{-\psi}{|\nabla \phi|} \,. \quad (6)$$

Any tangent vector can be obtained in this way.

Consider a surface functional $E : M \to \mathbb{R}$. When the function $E(\Gamma(s))$ is derivable at $s = 0$ we say that the *Gâteaux derivative* of E at Γ in the direction v is

$$\partial E(\Gamma, v) \equiv \frac{d}{ds} E(\Gamma(s))\Big|_{s=0} \,. \quad (7)$$

If this derivative can be written as

$$\partial E(\Gamma, v) = \int_\Gamma w(\mathbf{x}) \, v(\mathbf{x}) \, d\sigma \,, \quad (8)$$

with w in the tangent space, we say that w is the gradient of E. This allows to evolve the surface in the direction, $-w$, which ensures a decrease of the functional. Indeed, if $\Gamma(t)$ satisfies $\frac{\partial}{\partial t}\Gamma = -w$, then we have $\frac{\partial}{\partial t}E(\Gamma(t)) = -\int_\Gamma w \cdot w \, d\sigma \leq 0$. Thus, an evolution in the direction $-w$ will decrease the functional.

4. Understanding the Visibility

This section presents an analysis of the visibility and its evolution. The analysis stands on the study of Tsai *et al.* [22], who described the dynamics of the visible regions as the observer moves. The goal here, is to compute the derivative of the visibility function with respect to surface variations instead.

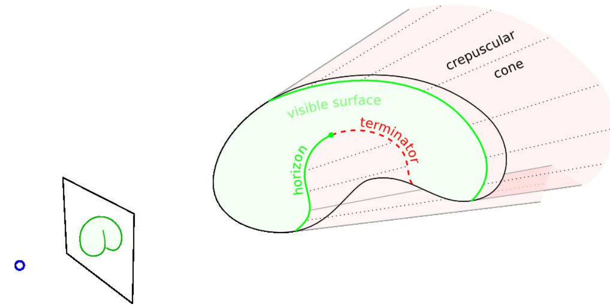

Figure 1. The *banana* shape seen from a vantage point. The horizon is drawn with green lines and the terminators with a dashed red line. The viewing ray segments are the crepuscular rays that form the crepuscular cone.

4.1. Geometrical Description

Given a solid shape and a vantage point, a point is said to be visible if no point of the shape lies on the segment between it and the vantage point. The set of all visible points will be called the *visible volume*, $\mathcal{V}$, and its complement, $\mathcal{V}^c$, the *occluded volume*. These volumes are respectively colored in green and red in Figure 2. The frontier between these two volumes, $\partial\mathcal{V}$, is a surface and will be called the *visibility interface*, see Figures 1 and 2.

The visibility interface is composed of two parts of different nature, one that is visible (the green curves in Figure 2) and one that is occluded (the red ones). The visible part of the interface coincide with the visible part of the shape's surface and background. Therefore, we will refer to this part as the *visible surface*. The occluded part of the interface is mostly in the free space. It is formed by patches of a generalized cone joining different parts of the visible surface. In analogy to atmospherical optics, we will call this part *crepuscular cone* (cf. also to Figure 1).

The border between the visible surface and the crepuscular cone is a closed curve on the shape's surface. Again, this curve contains both visible and occluded points. The visible part is the *horizon* or *contour generator* (green curves in Fig. 1). It is a (possibly open) curve made of visible surface points whose normal is perpendicular to the viewing ray. Its projection into the image are the *apparent contours*. The occluded part is the *terminator* (red curves in Fig. 1). It contains the points where the shadow of the horizon is cast. The segments joining points in horizon with their terminators are the *crepuscular rays* that form the crepuscular cone. The points in the crepuscular cone are all occluded by the horizon.

4.2. Mathematical Formulation

Let the visibility function $\nu_\Gamma : \mathbb{R}^3 \to \{0, 1\}$ be the characteristic function of the visible volume, the binary function evaluating to 1 for points that are visible and to 0 elsewhere.

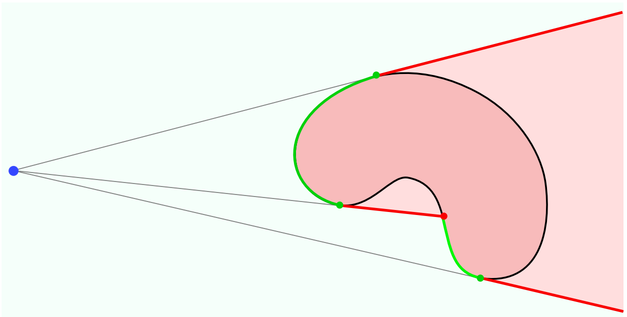

Figure 2. Sliced view of the banana shape. The visible surface/volume is drawn in green and the crepuscular cone/occluded volume in red.

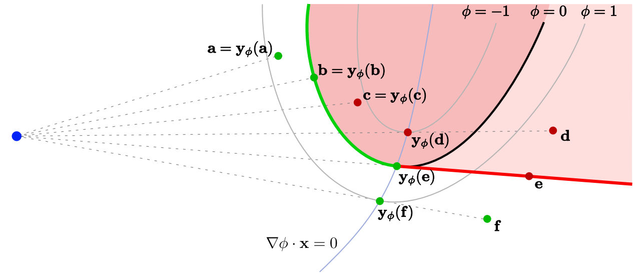

Figure 3. Location of $\mathbf{y}_\phi(\mathbf{x})$ according to the position of $\mathbf{x}$. If $\mathbf{x}$ is in the visible volume then $\phi(\mathbf{y}_\phi(\mathbf{x}))$ is positive, else it is negative.

In this section we write this function in terms of the level set function ϕ in order to derive analytical expressions for its spatial and temporal derivatives in the next sections. Later, the results will be used for computing the Gâteaux derivative of the reprojection error functional.

Assume the vantage point to be at the origin and let ϕ be an implicit representation of the surface. The visibility of a point $\mathbf{x}$ can be determined from the values that ϕ takes along the segment connecting the origin with $\mathbf{x}$. If any of these values is negative, then $\mathbf{x}$ is occluded.

Let $\mathbf{y}_\phi(\mathbf{x})$ be the point of the segment where ϕ admits the minimum (see Figure 3), i.e. $\mathbf{y}_\phi(\mathbf{x}) = \alpha_\phi(\mathbf{x})\mathbf{x}$ with

$$\alpha_\phi(\mathbf{x}) = \arg\min_{\alpha\in[0,1]} \phi(\alpha\mathbf{x}) \ . \tag{9}$$

If the minimum is not unique, take the closest to the origin. We observe that $\phi(\mathbf{y}_\phi(\mathbf{x}))$ is negative in the interior of the occluded volume and positive in the interior of the visible volume (cf. Figure 3). It can be shown that $\phi \circ \mathbf{y}_\phi$ is a continuous function [22], therefore, $\phi(\mathbf{y}_\phi(\mathbf{x})) = 0$ for all the points on the visibility interface regardless of their visibility.

This implies that for every point $\mathbf{x}$ on the visibility interface, $\mathbf{y}_\phi(\mathbf{x})$ is a point on the surface. If $\mathbf{x}$ is itself on the visible surface then necessarily $\mathbf{y}_\phi(\mathbf{x}) = \mathbf{x}$, otherwise $\mathbf{x}$ would be occluded. If $\mathbf{x}$ is on a crepuscular ray, then $\mathbf{y}_\phi(\mathbf{x})$ is its occluder, lying on the horizon where the crepuscular ray begins. All points on a crepuscular ray share the same occluder, $\mathbf{y}_\phi$. This fact gives an important role to the horizon and its consequences will be seen in the following sections.

From above, it follows that $\phi \circ \mathbf{y}_\phi$ is an implicit representation of the closure of the occluded volume and

$$\nu_\Gamma(\mathbf{x}) = H(\phi(\mathbf{y}_\phi(\mathbf{x}))) \tag{10}$$

almost everywhere, with exactly the exception of the visibility interface. Also, as distributions, $\nu_\Gamma = H \circ \phi \circ \mathbf{y}_\phi$.

4.3. Spatial Derivative of the Visibility

The gradient of the visibility, $\nabla\nu_\Gamma$, is a distribution that computes flow integral across the visibility interface $\int \nabla\nu_\Gamma \cdot \mathbf{w} d\mathbf{x} = \int_{\partial\mathcal{V}} \mathbf{w} \cdot d\boldsymbol{\sigma}$ (see section 3.1). It can be imagined as vector field that is zero everywhere except on the visibility interface, where it is aligned with the interface's normal and is infinitely long.

In order to derive an analytical expression of the gradient $\nabla\nu_\Gamma$, we first note that

$$\nabla(\phi \circ \mathbf{y}_\phi) = \nabla\phi(\mathbf{y}_\phi)\alpha_\phi \tag{11}$$

almost everywhere. Specifically, this holds for all the points on the visibility interface except for the terminators where $\phi \circ \mathbf{y}_\phi$ is not derivable. To see this, we distinguish two cases. If $\mathbf{y}_\phi$ is in the interior of the segment between vantage point and $\mathbf{x}$, then it is a local minimum of ϕ in the ray and thus $\nabla\phi(\mathbf{y}_\phi(\mathbf{x})) \cdot \mathbf{x} = 0$. As a consequence, the chain rule yields the above result. Otherwise, when $\mathbf{y}_\phi$ is at an extremum of the segment, we generally have that $\mathbf{y}_\phi(\mathbf{x}) = \mathbf{x}$ and $\alpha_\phi(\mathbf{x}) = 1$ in a neighborhood of $\mathbf{x}$ and so the same consequence holds.

Now, applying the chain rule to (10) it follows that

$$\nabla\nu_\Gamma = \delta(\phi(\mathbf{y}_\phi))\ \nabla\phi(\mathbf{y}_\phi)\ \alpha_\phi \ . \tag{12}$$

4.4. Temporal Derivative of the Visibility

Consider a variation $\phi^s = \phi + s\psi$ of ϕ and let $\Gamma(s)$ be the associated deformed surface. The visibility function $\nu_{\Gamma(s)}(\mathbf{x})$ is now a space-time function. Its derivative with respect to time is a scalar distribution concentrated on the visibility interface. It measures the variation of quantities integrated over the visible domain. Intuitively, that is the difference between the amount of mass that enters and that exits the visible volume as the surface evolves.

The chain rule gives

$$\frac{d}{ds}\nu_{\Gamma(s)}(\mathbf{x})\Big|_{s=0} = \delta(\phi(\mathbf{y}_\phi))(\psi(\mathbf{y}_\phi) + \nabla\phi(\mathbf{y}_\phi)\cdot\dot{\mathbf{y}}_\phi) \tag{13}$$

where $\dot{\mathbf{y}}_\phi$ is the temporal derivative of $\mathbf{y}_{\phi^s}$ at $s = 0$. If $\mathbf{y}_\phi$ is in the interior of the segment, then $\dot{\mathbf{y}}_\phi$ and $\mathbf{x}$ are collinear and orthogonal to $\nabla\phi(\mathbf{y}_\phi)$. Otherwise, if $\mathbf{y}_\phi = \mathbf{x}$ then $\dot{\mathbf{y}}_\phi = 0$. So, in any case, we have

$$\frac{d}{ds}\nu_{\Gamma(s)}(\mathbf{x})\Big|_{s=0} = \delta(\phi(\mathbf{y}_\phi))\,\psi(\mathbf{y}_\phi)\,. \tag{14}$$

4.5. Temporal Derivative of a Quantity Integrated over the Visible Volume

We have now the necessary tools to compute the Gâteaux derivative of a functional F that is the integral of a quantity f over the visible volume,

$$F(\Gamma) = \int_{\mathbb{R}^3} f(\mathbf{x})\nu_\Gamma(\mathbf{x})d\mathbf{x}\,. \tag{15}$$

This derivative will be used in the following section to easily derive the derivative of the reprojection error functional. The main difficulties are contained in this section.

From equation (14), we see that

$$\begin{aligned}\frac{d}{ds}F(\Gamma(s))\Big|_{s=0} &= \int_{\mathbb{R}^3} f(\mathbf{x})\frac{d}{ds}\nu_{\Gamma(s)}(\mathbf{x})\Big|_{s=0}d\mathbf{x}\\ &= \int_{\mathbb{R}^3} f(\mathbf{x})\psi(\mathbf{y}_\phi)\delta(\phi(\mathbf{y}_\phi))\,d\mathbf{x}\,,\end{aligned} \tag{16}$$

which we rewrite as an integral over the visibility interface

$$\frac{d}{ds}F(\Gamma(s))\Big|_{s=0} = \int_{\partial\mathcal{V}} f(\mathbf{x})\frac{\psi(\mathbf{y}_\phi)}{|\nabla(\phi\circ\mathbf{y}_\phi)|}\,d\sigma\,, \tag{17}$$

by noting that $d\sigma = |\nabla(\phi\circ\mathbf{y}_\phi)|\delta(\phi(\mathbf{y}_\phi))d\mathbf{x}$. Ideally, we would like to write the derivative as an integral over the surface Γ and not over the visibility interface (see section 3.2). With this aim, we split the integral into a sum of two integrals, one over each of the parts of the visibility interface.

(i) On the visible surface, we know that $\mathbf{y}_\phi = \mathbf{x}$ and, thus, the integral is simply

$$\int_{\Gamma\cap\mathcal{V}} f(\mathbf{x})\frac{\psi(\mathbf{x})}{|\nabla\phi(\mathbf{x})|}d\sigma\,. \tag{18}$$

Note that, without loss of generality, we can assume ψ to be null on the background and, therefore, the integral over the background surface vanishes.

(ii) On the crepuscular cone, $\mathbf{y}_\phi$ is the occluder of $\mathbf{x}$ and is a point on the horizon of the surface. All the points on a crepuscular ray share the same occluder on the horizon. The idea, here, is to attribute all the mass of a crepuscular ray to the origin of the ray on the horizon. This way, the integral over the crepuscular cone will be written as an integral over the horizon and, therefore, over the surface Γ.

Given an arc parametrization of the horizon, $\boldsymbol{\gamma} : I \to \mathbb{R}^3 : t \mapsto \boldsymbol{\gamma}(t)$, the crepuscular cone can be parameterized by $\mathbf{x}(r,t) = r\boldsymbol{\gamma}(t)$ with $t \in I$ and r in the interval $(1, T_{\boldsymbol{\gamma}(t)})$; where for any fixed t, $\mathbf{x}(r,t)$ covers the crepuscular ray from the horizon $\boldsymbol{\gamma}(t)$ to the associated terminator. By using this parametrization, equation (11) and the fact that $\mathbf{y}_\phi(\mathbf{x}(r,t)) = \boldsymbol{\gamma}(t)$, the surface integral (17) over the crepuscular cone is written as

$$\int_I\int_1^{T_{\boldsymbol{\gamma}(t)}} f(r\boldsymbol{\gamma}(t))\frac{\psi(\boldsymbol{\gamma}(t))}{|\nabla\phi(\boldsymbol{\gamma}(t))|}r^2|\boldsymbol{\gamma}(t)\times\boldsymbol{\gamma}'(t)|\,dr\,dt\,. \tag{19}$$

The terms depending on r can be gathered together into $L(\mathbf{x}) = \int_1^{T_\mathbf{x}} f(r\mathbf{x})r^2dr$, which cumulates the mass of f along the crepuscular rays. Also, let $\boldsymbol{\eta}(t)$ denote the normal vector to the horizon, that is tangent to the surface and points away from the observer. The integral is, then,[2]

$$\int_I L(\boldsymbol{\gamma}(t))\frac{\psi(\boldsymbol{\gamma}(t))}{|\nabla\phi(\boldsymbol{\gamma}(t))|}\,\left(\boldsymbol{\gamma}(t)\cdot\boldsymbol{\eta}(t)\right)\,dt\,. \tag{20}$$

This is the flux of the vector field $L\frac{\psi}{|\nabla\phi|}\mathbf{x}$ crossing the horizon from the visible surface to the crepuscular cone. It is an integral over the horizon. Let us now convert it into an integral over the surface Γ.

Let $\mathcal{O} = \{\mathbf{x}\cdot\mathbf{n}(\mathbf{x}) \le 0 : \mathbf{x}\in\Gamma\}$ be the set surface points whose normal is oriented towards the camera. The border of this set as a subset of the surface, $\partial\mathcal{O}$, is a curve on the surface. The horizon corresponds exactly to the visible part of this curve. From (4) on the Riemanian manifold Γ (instead of $\mathbb{R}^3$), we know that

$$\int_\Gamma \nabla_\Gamma\chi_\mathcal{O}\cdot\mathbf{w}\,d\sigma = -\int_{\partial\mathcal{O}}\mathbf{w}\cdot\boldsymbol{\eta}\,d\tau\,, \tag{21}$$

where ∇_Γ denotes the intrinsic gradient in Γ and $\chi_\mathcal{O} = 1-H(\mathbf{x}\cdot\mathbf{n})$ is the characteristic function of $\mathcal{O}$. Considering the vector field $\mathbf{w}(\mathbf{x}) = \nu_\Gamma(\mathbf{x})L(\mathbf{x})\frac{\psi(\mathbf{x})}{|\nabla\phi(\mathbf{x})|}\mathbf{x}$, which is zero for all the points of $\partial\mathcal{O}$ except the horizon, we have that equation (20) can be written as

$$\int_\Gamma \nu_\Gamma(\mathbf{x})L(\mathbf{x})\frac{\psi(\mathbf{x})}{|\nabla\phi(\mathbf{x})|}\,\left(\mathbf{x}\cdot\nabla_\Gamma[H(\mathbf{x}\cdot\mathbf{n})]\right)d\sigma\,. \tag{22}$$

Since $\nabla[H(\mathbf{x}\cdot\mathbf{n})]$ is on the tangent plane of Γ, it corresponds to $\nabla_\Gamma[H(\mathbf{x}\cdot\mathbf{n})]$. Thus, by (5), we can rewrite $\mathbf{x}\cdot\nabla_\Gamma[H(\mathbf{x}\cdot\mathbf{n})]$ as $\mathbf{x}^t\nabla\mathbf{n}\mathbf{x}\;\delta(\mathbf{x}\cdot\mathbf{n})$. Finally, joining the splited integrals (18) and (22), the Gâteaux derivative of F is

$$\partial F(\Gamma, v) = \int_\Gamma -\left(f + L\,\mathbf{x}^t\nabla\mathbf{n}\mathbf{x}\,\delta(\mathbf{x}\cdot\mathbf{n})\right)\nu_\Gamma\;v\;d\sigma. \tag{23}$$

Remark: As expected in section 3.2, we have managed to rewrite the Gâteaux derivative of F as $\int_\Gamma w\;v\;d\sigma$. Nevertheless, unusually here, w is not a function, but a distribution. Distributions are linear continuous operators, so the functional is Fréchet differentiable and the differential is w. However, the gradient in the tangent space, as defined in

[2] $\boldsymbol{\gamma}'$ and $\boldsymbol{\eta}$ form an orthonormal basis of the tangent plane at $\boldsymbol{\gamma}$. In this basis, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}\cdot\boldsymbol{\gamma}')\boldsymbol{\gamma}' + (\boldsymbol{\gamma}\cdot\boldsymbol{\eta})\boldsymbol{\eta}$ and thus $|\boldsymbol{\gamma}\times\boldsymbol{\gamma}'| = \boldsymbol{\gamma}\cdot\boldsymbol{\eta}$.

[20] and [3], does not exist for this functional, because w is not an admissible deformation of Γ. In other words, to perform a gradient descent evolution w has to be approximated by an admissible deformation. In practice, this reduces simply to approximating the delta distribution with a function.

5. Differential of the Reprojection Error

Using the results of the previous section, in this section we are going to compute the differential of the reprojection error functional (3),

$$E(\Gamma) = -\int_{(\Gamma \cup B) \cap \mathcal{V}} g(\mathbf{x}, \mathbf{n}(\mathbf{x})) \, \frac{\mathbf{x} \cdot \mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3} \, d\sigma \,. \quad (24)$$

Since, for all the points $\mathbf{x}$ of the visible surface, the normal $\mathbf{n}_{\partial\mathcal{V}}(\mathbf{x})$ to the visibility interface $\partial\mathcal{V}$ coincides with the normal $\mathbf{n}(\mathbf{x})$ to the surface Γ, and since $\mathbf{x} \cdot \mathbf{n}_{\partial\mathcal{V}}(\mathbf{x}) = 0$ on the crepuscular cone, it follows that the integral can be extended to the whole visibility interface

$$E(\Gamma) = -\int_{\partial\mathcal{V}} g(\mathbf{x}, \mathbf{n}(\mathbf{x})) \, \frac{\mathbf{x}}{\mathbf{x}_z^3} \cdot \mathbf{n}_{\partial\mathcal{V}}(\mathbf{x}) \, d\sigma \,. \quad (25)$$

5.1. Case where g does not depend on the Normal

Let us first consider the case where g does not depend on the normal, $g(\mathbf{x}, \mathbf{n}(\mathbf{x})) = g(\mathbf{x})$. The functional (25) is the flux of the vector field $g(\mathbf{x}) \frac{\mathbf{x}}{\mathbf{x}_z^3}$ across the visibility interface $\partial\mathcal{V}$. By Gauss' divergence theorem, this flux is the opposite of the amount of divergence of the vector field inside the visible volume. Thus, as $\nabla \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3} = 0$,

$$E(\Gamma) = \int_{\mathbb{R}^3} \left(\nabla g(\mathbf{x}) \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3} \right) \nu_\Gamma(\mathbf{x}) \, d\mathbf{x} \,. \quad (26)$$

Now, by using the result (23) developed in the previous section with $f(\mathbf{x}) = \nabla g(\mathbf{x}) \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3}$, we immediately get the Gâteaux derivative of this functional. We observe that, in this case, L has a simple form, because the integral sums up the variations of g along the crepuscular rays. This is

$$L(\mathbf{x}) = \int_1^{T_\mathbf{x}} \nabla g(r\mathbf{x}) \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3} dr = \big[g(T(\mathbf{x})) - g(\mathbf{x}) \big] \frac{1}{\mathbf{x}_z^3} \,, \quad (27)$$

where $T(\mathbf{x})$ is the terminator of $\mathbf{x}$. Finally, noting $g \circ T$ by g', the differential of the reprojection error functional is

$$-\nabla g \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3} \nu_\Gamma + (g - g') \frac{\mathbf{x}^t \nabla \mathbf{n} \mathbf{x}}{\mathbf{x}_z^3} \delta(\mathbf{x} \cdot \mathbf{n}) \nu_\Gamma. \quad (28)$$

5.2. With Normals

We describe here the derivation of the differential for the general case where g may depend on the normal of the surface. Because of space limitation, we will only sketch the calculus and present the final result.

First, using (4) with the visibility interface $\partial\mathcal{V}$ (instead of the surface Γ), we reformulate (25) as an integral over $\mathbb{R}^3$. Then, deriving the result with respect to s, the product rule yields

$$\begin{aligned} \frac{d}{ds} E(\Gamma(s)) \Big|_{s=0} = &- \int_{\mathbb{R}^3} \frac{d}{ds} g(\mathbf{x}, \mathbf{n}^s) \Big|_{s=0} \frac{\mathbf{x}}{\mathbf{x}_z^3} \cdot \nabla \nu_\Gamma \, d\mathbf{x} \\ &- \int_{\mathbb{R}^3} g(\mathbf{x}, \mathbf{n}) \frac{\mathbf{x}}{\mathbf{x}_z^3} \cdot \nabla \frac{d}{ds} \nu_{\Gamma(s)} \Big|_{s=0} \, d\mathbf{x} \,. \end{aligned}$$

For the first integral, we have $\frac{d}{ds} g(\mathbf{x}, \mathbf{n}^s)|_{s=0} = g_\mathbf{n} \cdot \frac{\nabla \psi}{|\nabla \phi|}$. To get rid of the $\nabla \psi$ term and make ψ appear instead, one has to do integration by parts on $\nabla \psi$, as done in [20], section 5. The second integral, also by integration by parts, becomes $-\int_{\mathbb{R}^3} \nabla \cdot [g(\mathbf{x}, \mathbf{n}) \frac{\mathbf{x}}{\mathbf{x}_z^3}] \frac{d}{ds} \nu_{\Gamma(s)} \big|_{s=0} d\mathbf{x}$ and so, as $\mathbf{n}$ does not depend on s, one can apply the result (28) of the simpler case when g does not depend on the normal.

The resulting differential is

$$-\nabla \cdot \left(g_\mathbf{n} \frac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3} + g \frac{\mathbf{x}}{\mathbf{x}_z^3} \right) \nu_\Gamma + (g - g') \frac{\mathbf{x}^t \nabla \mathbf{n} \, \mathbf{x}}{\mathbf{x}_z^3} \delta(\mathbf{x} \cdot \mathbf{n}) \nu_\Gamma \,. \quad (29)$$

5.3. Comparison with the Weighted Area

If we denote $\overline{g}(\mathbf{x}, \mathbf{n}) = g(\mathbf{x}, \mathbf{n}) \frac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3}$, then $[\overline{g}_\mathbf{n} + \overline{g} \; \mathbf{n}]$ corresponds to $[g_\mathbf{n} \frac{\mathbf{x} \cdot \mathbf{n}}{\mathbf{x}_z^3} + g \frac{\mathbf{x}}{\mathbf{x}_z^3}]$. So the differential (29) is equal to the gradient of the weighted area functional $\int_\Gamma \overline{g}(\mathbf{x}, \mathbf{n}(\mathbf{x})) \, d\sigma$ given in [7, 20], plus a new term $(g - g') \frac{\mathbf{x}^t \nabla \mathbf{n} \, \mathbf{x}}{\mathbf{x}_z^3} \delta(\mathbf{x} \cdot \mathbf{n}) \nu_\Gamma$ which is due to the changes of visibility caused by the movement of the horizon.

6. Application to Multi-view Stereo

In this section we present a sample application of the reprojection error functional to multi-view stereo. To keep the example simple, the scene is assumed to be Lambertian and the illumination static. Note though that more elaborate reflectance models still lead to a reprojection error functional.

The image formation model is the same as the one of stereoscopic segmentation [25]. To explain the images, one needs a surface Γ and also the radiance of points of that surface and of the background. Let $C : \mathbb{R}^3 \to \mathbb{R}^3$ be the radiance function that associates colors to the points of the 3D space and the background. Ideally, the color $I(\mathbf{u})$ observed at the pixel $\mathbf{u}$ of image I should be equal to the color of its backprojection onto the surface $C(\pi_\Gamma^{-1}(\mathbf{u}))$. Thus, the SSD reprojection error of the surface into an image is

$$E(\Gamma, C) = \int_\mathcal{I} \big(I(\mathbf{u}) - C(\pi_\Gamma^{-1}(\mathbf{u})) \big)^2 \, d\mathbf{u} \,. \quad (30)$$

The reprojection error for a set of images is the sum of the individual reprojection errors. An additional smoothing area energy $\int_\Gamma d\sigma$ is also added to represent our prior belief that surfaces are smooth.

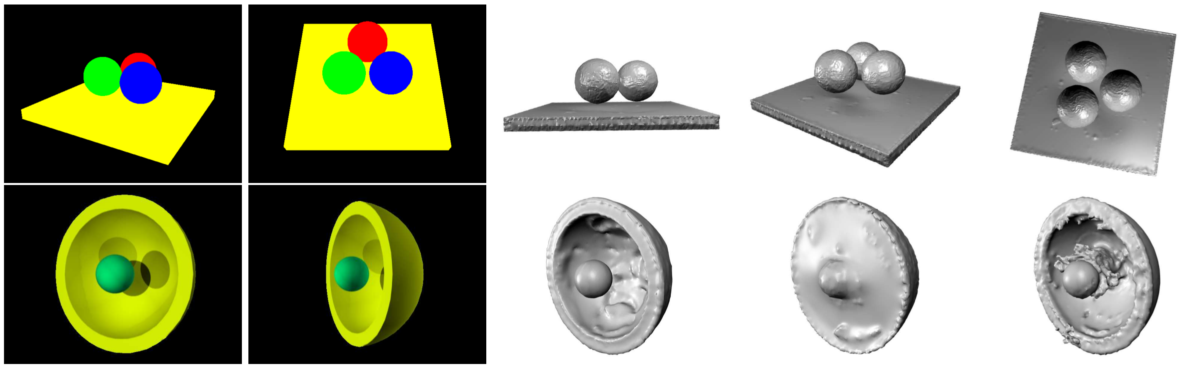

Figure 4. **Top:** two input images of the *balls* dataset and three rendereings of the reconstruction obtained with the horizon term. **Bottom:** two input images of the *bowl* dataset and the reconstruction obtained by the full term, the horizon term and the interior term.

The optimization is done by alternating between the estimation of C and Γ. For a fixed surface, the optimal radiance of a point has a closed form solution as a weighted sum of the colors observed at its projection onto the images where it is visible. For a fixed radiance, from equation (29), the differential of a single image error with respect to the surface is

$$(I-C)^t \nabla C \frac{\mathbf{x}}{\mathbf{x}_z^3} \nu_\Gamma + \left((I-C)^2 - (I-C')^2\right) \frac{\mathbf{x}^t \nabla \mathbf{n} \mathbf{x}}{\mathbf{x}_z^3} \delta(\mathbf{x} \cdot \mathbf{n}) \nu_\Gamma \tag{31}$$

where C' denotes the radiance at the terminator of $\mathbf{x}$.

Intuitively, this means that during the evolution, the visible points will move according to the first term of (31) in order to match C with I in the interior of objects in the image. Additionally, the second term will move the horizon of the surface and only the horizon because of the $\delta(\mathbf{x} \cdot \mathbf{n})\nu_\Gamma$ factor. This term compares the cost of the points in the horizon with the cost of the terminator and moves the horizon accordingly, so that the terminator becomes visible or occluded depending on that comparison. As a consequence, the apparent contours of the surface on the image will move to their correct location, as will be shown in the experiments.

6.1. Experiments

We implemented the surface evolution of (31) using the level set method in a multi-resolution scheme. Visibility is computed by rendering the surface using graphics hardware and then comparing the depth of the points with the Z-buffer. The horizons are found by approximating the delta distribution with a Gaussian. Radiance at terminators is computed by re-rendering the surface, horizons excluded.

We present here, the experiments performed on two, specially designed, synthetic scenes and two real world scenes. The goal of these experiments is to show the impact of the proper handling of the visibility, not to evaluate the performance of the generative model (30) presented above.

The *balls* dataset (fig. 4) consists of 20 images of three balls floating above a plane. There is no texture or shading in any part of the scene. Therefore, the only information present in the images are the apparent contours. In addition, because of self-occlusions between the balls and the plane, the silhouettes of the foreground are not sufficient to distinguish that the balls are three separate objects.

The reprojection error minimizing flow (31) was executed 3 times. First, using the flow as it is, then, using only its second term (the horizon term) and, finally, using only the first term (the interior term). The first two executions successfully managed to separate the three balls and obtained a correct reconstruction. The third one, did not separate the balls during the evolution and, due to the lack of texture, did shrink and disappear. The shrinkage did happen even when initializing from the ground truth.

We repeated the experiment for the *bowl* scene (fig. 4). The scene contains a green ball inside a yellow bowl with Lambertian shading. The execution with the full flow, correctly recovered the concavity of the bowl and the shape of the ball. The execution using only the horizon term did not carve the concavity at all. The execution with the interior term, did carve the concavity, but not completely, keeping the ball and the bowl linked together. This shows how the interior and the horizon terms worked together, the first one carving the concavity and the second one enforcing the apparent contour of the ball on the images.

Finally, we tested the evolution on the dino and temple datasets of the multi-view stereo database [17]. Numerical evaluation of the reconstructions performed by D. Scharstein and B. Curless can be found at `http://vision.middlebury.edu/mview/`. While the final reconstructions are not specially precise, probably due to the simplicity of the model (30), it is interesting to see the evolution itself (Figure 5). Two initial ellipsoids deform and *grow* according to the horizon term until fully explaining the input images. The evolution took 40 mins, the last 30 mins of which the surface reminded nearly steady.

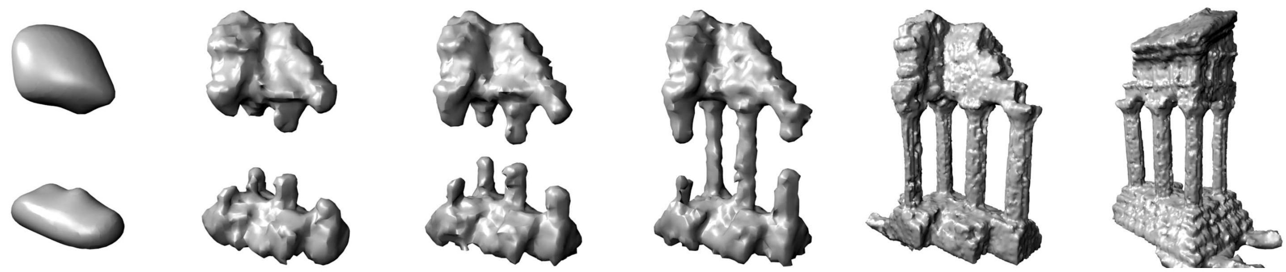

Figure 5. Surface evolution for the *temple sparse ring* dataset and the final reconstruction.

7. Conclusion

In this paper we compute the derivative of the reprojection error functional. The difficult part has been to correctly take into account the visibility changes that occur while the surface moves, which is one of the most challenging problems in surface reconstruction from images. The reward, is that it is now possible to minimize the reprojection error via surface evolution.

The benefit of this minimization is that the reconstructed surface is the one that best reproduces the observed images. In particular, as demonstrated in the experiments, the evolution moves the contour generators of the surface so that the apparent contours appear at their correct location in the images. This is a direct consequence of the correct minimization of the reprojection error itself. Therefore, current methods using additional silhouettes or apparent contour constraints can now be understood and justified by a single criterion: the reprojection error.

Acknowledgments The authors wish to thank Jean-Philippe Pons for sharing his level set method implementation and S.M. Seitz, B. Curless, J. Diebel, D. Scharstein and R. Szeliski for the temple and dino datasets and evaluations.

References

[1] B. Appleton, H. Talbot. Globally minimal surfaces by continuous maximal flows. *PAMI*, 28(1):106-118, 2006. 1, 2

[2] Y. Boykov, V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. *ICCV*, 2003. 1

[3] G. Charpiat, R. Keriven, J. P. Pons, O. Faugeras. Designing spatially coherent minimizing flows for variational problems based on active contours. *ICCV*, 2:1403–1408, 2005. 6

[4] D. Crispell, D. Lanman, P.G. Sibley, Y. Zhao, G. Taubin. Beyond Silhouettes: Surface Reconstruction using Multi-Flash Photography. *3DPVT*, 2006. 2

[5] O. Faugeras, R. Keriven. Complete dense stereovision using level set methods. *ECCV*, I:379–393, 1998. 1, 2

[6] Y. Furukawa, J. Ponce. Carved visual hulls for image-based modeling. *ECCV*, 2006. 2

[7] B. Goldlücke, I. Ihrke, C. Linz, M. Magnor. Weighted minimal hypersurface reconstruction. *PAMI*, 29(7):1194–1208, 2007. 1, 6

[8] C. Hernández Est., F. Schmitt. Silhouette and stereo fusion for 3D object modeling. *CVIU*, 96(3):367–392, 2004. 2

[9] R. Keriven. A variational framework for shape from contours. Tech. Report, CERMICS, ENPC, 2002-221b. 2

[10] J. Koenderink. *Solid shape*. MIT Press, 1990. 2

[11] V. Lempitsky, Y. Boykov, D. Ivanov. Oriented visibility for multiview reconstruction. *ECCV*, III:226–238, 2006. 1, 2

[12] S. Osher, J.A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. of Comput. Physics*, 79:12–49, 1988. 2

[13] S.J. Osher, R.P. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 2002. 2

[14] S. Paris, F.X. Sillion, L. Quan. A surface reconstruction method using global graph cut optimization. *IJCV*, 66(2):141–161, 2006. 2

[15] J.P. Pons, R. Keriven, O. Faugeras. Modelling dynamic scenes by registering multi-view image sequences. *CVPR*, 2:822–827, 2005. 2

[16] W. Rudin. *Functional Analysis*. McGraw-Hill, 1991. 3

[17] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. *CVPR*, 1:519–528, 2006. 2, 7

[18] S.N. Sinha, M. Pollefeys. Multi-view reconstruction using photo-consistency and exact silhouette constraints: a maximum-flow formulation. *ICCV*, 1:349–356, 2005. 2

[19] S. Soatto, A.J. Yezzi, H. Jin. Tales of shape and radiance in multi-view stereo. *ICCV*, 974–981, 2003. 2

[20] J.E. Solem, N. Overgaard. A geometric formulation of gradient descent for variational problems with moving surfaces. *Scale-Space*, 419–430, 2005. 1, 2, 3, 6

[21] C. Strecha, R. Fransens, L. Van Gool. Wide-baseline stereo from multiple views: a probabilistic account. *CVPR* 2004. 2

[22] Y.H.R. Tsai, L.T. Cheng, S. Osher, P. Burchard, G. Sapiro. Visibility and its dynamics in a PDE based implicit framework. *J. of Comput. Physics*, 199(1):260–290, 2004. 3, 4

[23] G. Vogiatzis, P. Torr, R. Cipolla. Multi-view stereo via volumetric graph-cuts. *CVPR*, 2:391–398, 2005. 2

[24] R.T. Whitaker. A Level-Set Approach to 3D Reconstruction from Range Data. *IJCV*, 29(3):203–231, 1998. 1

[25] A. Yezzi, S. Soatto. Stereoscopic segmentation. *IJCV*, 53(1):31–43, 2003. 2, 6

[26] T. Yu, N. Ahuja, W.-C. Chen. SDG cut: 3D reconstruction of non-lambertian objects using graph cuts on surface distance grid. *CVPR*, 2:2269–2276, 2006. 2